Supplementary Material

1. Hyperparameter Analysis

The information constraint I_c is an important hyperparameter in SIBAN to control the feature purification. Here we train SIBAN with varying I_c over a range {100, 200, 300}, and report the corresponding KL-divergence curves for both domains in Fig. 1, as well as the corresponding adaptive β_S/β_T values over the training course in Fig. 2. From the results, we can see that the information can be easily constrained to a specific I_c when choosing a relatively large I_c between $[200, 300]$. When choosing $I_c = 100$, the feature purification process becomes harder, since an information bottleneck with such a small I_c (100) is too narrow to maintain the necessary information for segmentation (see Fig. 1). Accordingly, the model has to give more bias to decrease the information constraint loss when choosing a small I_c, which explains why the β_S/β_T are relative large during the training course (see Fig. 2). We also observe that the information from the target domain is easier to be constrained to I_c. This is because the target model is trained under an unsupervised mechanism, which is more easily dominated by the information constraint loss.

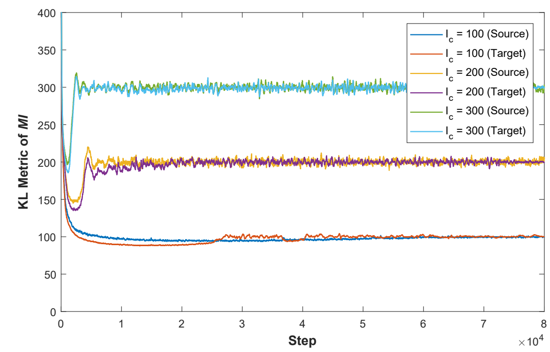

Figure 1: KL-divergence curves over the course of training.

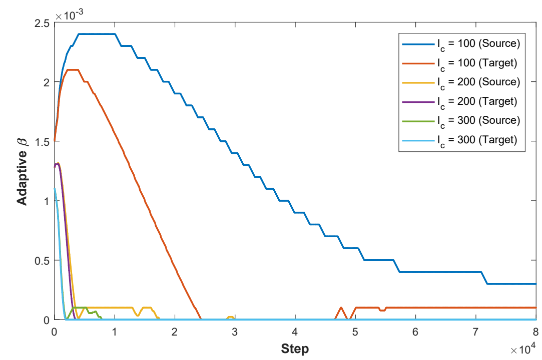

Figure 2: The values of adaptive β_S/β_T over the course of training.

2. Feature Distribution Visualization

In this section, we visualize the feature distributions in latent space aiming to confirm the effects of our method. To this end, we first select two similar images (x_S and x_T) from source and target domain respectively and then map their high-dimensional latent features (z_S and z_T) to a 2-D space with t-SNE shown in Fig. 3. In the first row, we label the t-SNE maps by different domains in order to evaluate the marginal distribution alignment (**global alignment**) of the features between domains. While in the second row, we label the t-SNE maps by different semantic classes in order to evaluate the semantic consistency (**local alignment**) of the features between domains.

From the t-SNE maps, we can observe that the non-adaptive model can not yield well-aligned latent features, neither in global level (see Fig. 3a) or class level (see Fig. 3d). These results demonstrate that the classifier trained on source data cannot be directly applied to target samples due to its limit generalization ability. For IBAN, the marginal distributions of the two domains are well aligned (see Fig. 3b), but some features from different semantic classes are mismatched (see Fig. 3e). The reason lies in that the information constraint in IBAN, which enforces task-dependent features to a standard Gaussian distribution, would wrongly compress the features from different classes too close to others and therefore make it hard to align them. Finally, we can see that SIBAN achieves good global and local feature alignment between domains (see Fig. 3c & Fig. 3f). The visualization of the latent feature distributions further explains why the SIBAN can achieve the leading results in feature-space adaptation.

3. More Qualitative Results

In Fig. 4, we show more qualitative results from the baseline method, IBAN, and SIBAN respectively, followed by the ground truth label map.

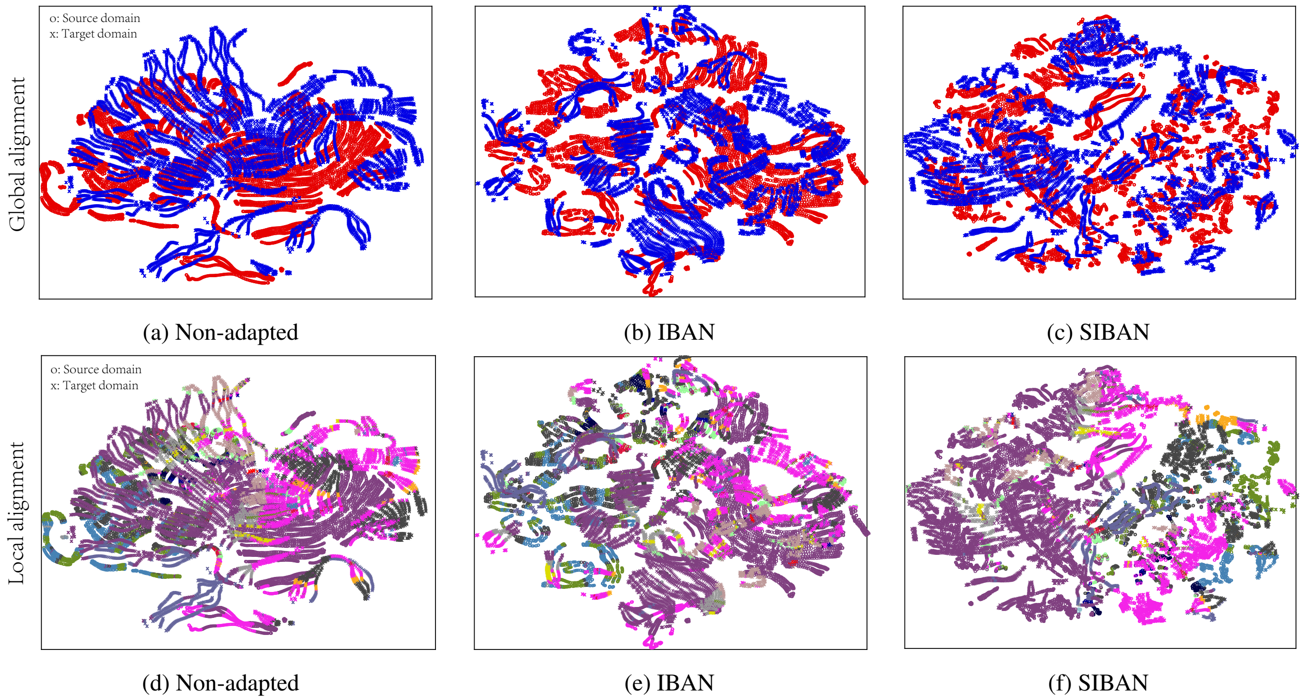

Figure 3: (Better zoom in.) We confirm the effects of SIBAN through a visualization of the learned representations z_S & z_T using t-distributed stochastic neighbor embedding (t-SNE). Specifically, we show the results of Non-adapted model in (a)&(d), IBAN in (b)&(e) and SIBAN in (c)&(f), respectively. In the first row, we label the t-SNE map by domains, where red denotes the source domain and blue denotes the target domain. In the second row, we label the t-SNE map by different classes. The colors are consistent with the annotation maps.

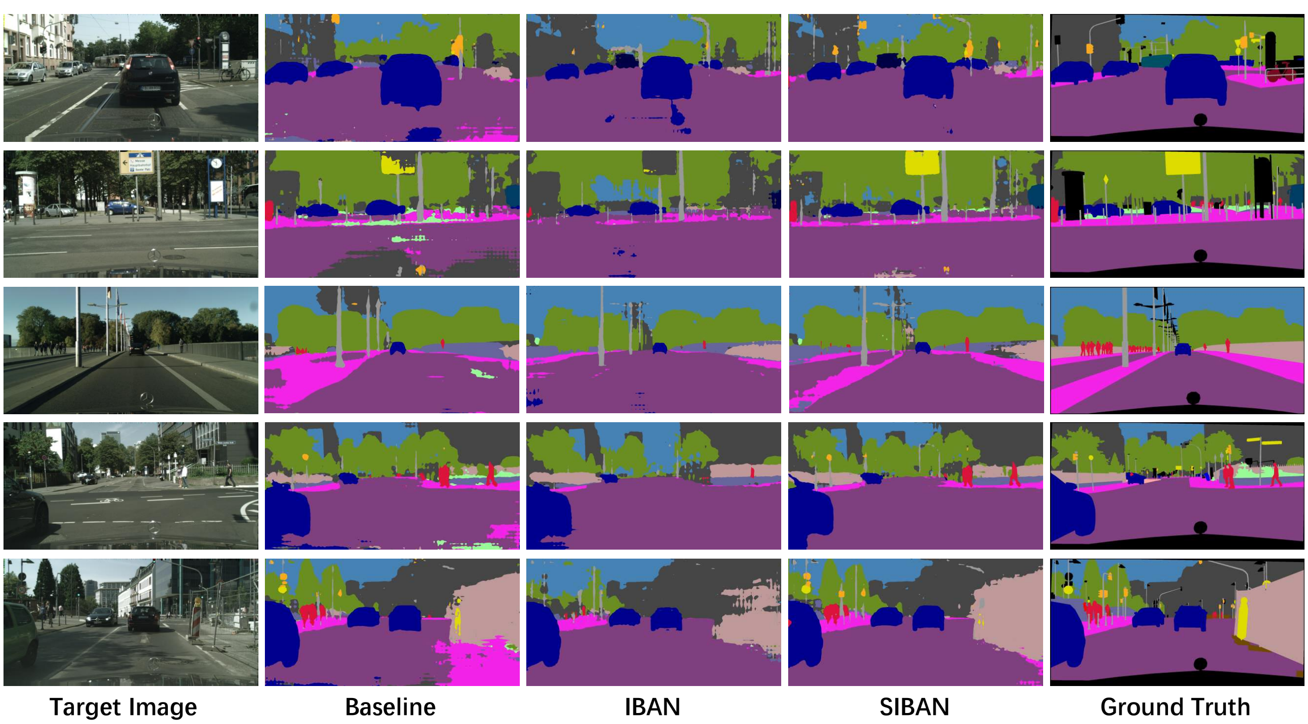

Figure 4: Qualitative results of the domain adaptive segmentation.